REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 08/26/2005	3. REPORT TYPE AND DATES COVERED Final Progress Report 01 NOV 04 - 31 JUL 05

4. TITLE AND SUBTITLE Sensor Modeling and Multi-Sensor Data Fusion	5. FUNDING NUMBERS W911NF-04-1-0434
6. AUTHOR(S) Devendra P. Garg and Manish Kumar	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Department of Mechanical Engineering and Materials Science Duke University, Durham, NC 27708	8. PERFORMING ORGANIZATION REPORT NUMBER RAMA-ARO-101
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSORING / MONITORING AGENCY REPORT NUMBER 47513.1-CI-11

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.	12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This research report presents a novel strategy to develop a sensor model based on a probabilistic approach that would accurately provide information about individual sensor's uncertainties and limitations. The strategy also establishes the dependence of sensor's uncertainties on some of environmental parameters or parameters of any feature extraction algorithm used in estimation based on sensor's outputs. The approach makes use of a neural network that is trained with the help of an innovative technique that obtains training signal from a maximum likelihood estimator. The proposed technique was applied for modeling stereo-vision sensors and an Infra-Red (IR) proximity sensor used in the robotic workcell available in the Robotics and Manufacturing Automation (RAMA) Laboratory at Duke University. In addition, the report presents an innovative method to fuse the probabilistic information obtained from these sensors based on Bayesian formalism in an occupancy grid framework to obtain three-dimensional occupancy model and key features of the robotic workspace. The capability of the proposed technique in accurately obtaining three-dimensional occupancy profile and efficiently removing individual sensor uncertainties was validated and compared with other methods via experiments carried out in the RAMA lab during this project.

14. SUBJECT TERMS Sensor Modeling, Sensor Fusion, Occupancy Grid, Maximum Likelihood, Neural Network	15. NUMBER OF PAGES 18
	16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	**UNCLASSIFIED**	**UNCLASSIFIED**	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

Enclosure 1

TABLE OF CONTENTS

Sensor Modeling and Multi-Sensor Data Fusion

1. INTRODUCTION

The desire to enhance the capabilities of modern control systems has led to the development of complex systems that are characterized by their increased non-linearity, flexibility, intelligence and enhanced ability to handle uncertainty. In order to incorporate autonomous decision making abilities, these systems need to possess complex capabilities such as perception, knowledge acquisition, learning, adaptability, and reasoning. Moreover, the system should be able to draw inference from incomplete, ambiguous or approximate information, and deal with uncertain and dynamic situations. The revolutionary advancement in the field of sensor technology that has led to development of superior sensing capabilities, and progress in computing and information processing has made it possible to develop systems with autonomous decision making capabilities.

Dynamic systems generally employ multiple sensors to provide diverse, complementary as well as redundant information. The primary goal of a multi-sensor system is to combine information from a multitude of sources into a robust, accurate and consistent environment description. There are several issues that arise while solving the multiple sensor fusion problem including inherent uncertainty in each sensor's measurements, and diverse and temporally or spatially disparate nature of measurements. The uncertainties in sensors not only arise from the impreciseness and noise in the measurements, but are also caused by the ambiguities and inconsistencies in the environment, and inability to distinguish between them. The strategies used to fuse data from these sensors should be able to eliminate such uncertainties, take into account the environmental parameters that affect sensor measurements, and fuse different nature of information to obtain a consistent description of the environment.

The algorithms reported in literature to fuse data from multiple sensory sources can be classified into three categories: 1) Fusion based on probabilistic methods, 2) Fusion based on least-squares techniques, and 3) Fusion based on intelligent methods. All of these methods differ in the manner they try to model the uncertainties inherent in the sensor measurements. The research carried out in this project combines two different kinds of sensing modalities to obtain three-dimensional occupancy profiles of a robotic workspace. The first modality is two vision sensors mounted on a stereo rig, and the second one is an Infra-Red (IR) proximity sensor. The research presented in this report first discusses a neural network based novel technique to obtain probabilistic sensor models. The fusion is carried out in an occupancy grid framework with the help of Bayesian approach. Some of the other techniques used in literature for sensor fusion include Dempster Shafer theory for evidential reasoning [1-2], fuzzy logic [3-4], and statistical techniques [5] such as Kalman filter [6-8]. The following section briefly discusses the Bayesian approach.

2. BAYESIAN APPROACH FOR SENSOR FUSION

Bayesian inference [9-11] is a statistical data fusion algorithm based on Bayes' theorem [12] of conditional or *a posteriori* probability to estimate an n-dimensional state vector 'X', after the observation or measurement denoted by 'Z' has been made. The probabilistic information contained in Z about X is described by a probability density function (p.d.f.) $p(Z \mid X)$, known as likelihood function, or the sensor model, which is a sensor dependent objective function based on observation. The likelihood function relates the extent to which the *a posteriori* probability is subject to change, and is evaluated either via offline experiments or by utilizing the available information about the problem. If the information about the state X is made available independently before any observation is made, then likelihood function can be improved to provide more accurate results. Such *a priori* information about X can be encapsulated as the prior probability $P(X = x)$ and is regarded as subjective because it is not based on observed data. Bayes' theorem provides the posterior conditional distribution of $X = x$, given $Z = z$, as

$$p(X = x \mid Z = z) = \frac{p(Z = z \mid X = x)P(X = x)}{\int p(Z = z \mid X = x)P(X = x)dx} = \frac{p(Z = z \mid X = x)P(X = x)}{P(Z = z)} \tag{1}$$

Since the denominator depends only on the measurement (the summation is carried out over all possible values of state), an intuitive approach to the estimation can be made by maximizing this posterior distribution, i.e., by maximizing the numerator of (1). This is called *Maximum a posteriori* (or MAP) estimate, and is given by:

$$\hat{x}_{MAP} = \arg\max p(X = x \mid Z = z) \propto p(Z = z \mid X = x)P(X = x) \tag{2}$$

Another popular estimation scheme (called *Minimum Mean Square Error* (MMSE) estimator) minimizes the sum of square of errors, i.e., minimizes the Euclidean distance between the true state and the estimate after the observation has been made.

To incorporate the measurements from two sensors, (1) can be written as:

$$p(X = x \mid Z = z_1, z_2) = \frac{p(Z = z_1 \mid X = x)p(Z = z_2 \mid X = x)P(X = x)}{P(Z = z_1, z_2)} \tag{3}$$

Fig. 1 shows a process in which the sensory data made available from multiple sensors can be fused under the Bayesian scheme.

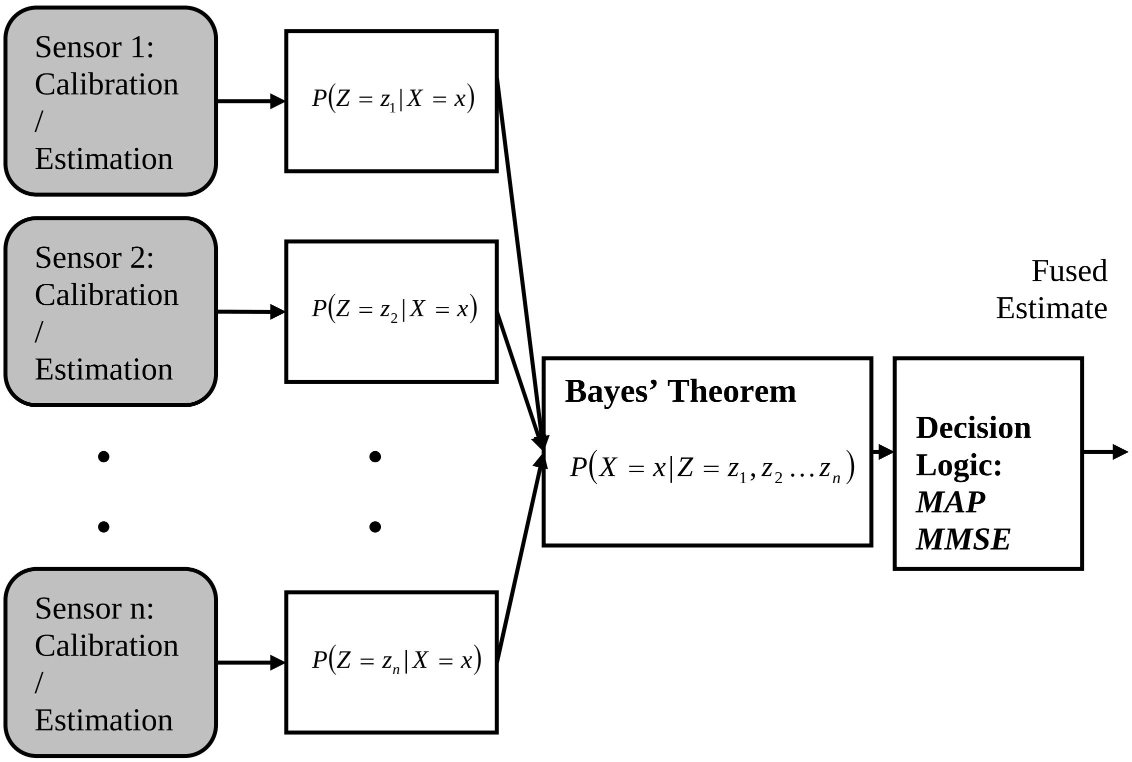

Fig. 1. Multiple Sensor Fusion using Bayesian Technique

3. SENSOR MODELS

Sensor modeling [13] deals with developing an understanding of the nature of measurements provided by the sensor, the limitations of the sensor, and probabilistic understanding of the sensor performance in terms of the uncertainties. The information supplied by a sensor is usually modeled as a mean about a true value, with uncertainty due to noise represented by a variance that depends on both the measured quantities themselves and the operational parameters of the sensor. A probabilistic sensor model is particularly useful because it facilitates the determination of the statistical characteristics of the data obtained. This probabilistic model is usually expressed in the form of probability density function (p.d.f.) $p(z|x)$ that captures the probability distribution of measurement by the sensor (z) when the state of the measured quantity (x) is known. This distribution is extremely sensor specific and can be experimentally determined. Gaussian distribution is one of the most commonly used distributions to represent the sensor uncertainties and is given by the following equation:

$$p(z|x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left\{\frac{-(x-z)^2}{2\sigma^2}\right\}} \tag{4}$$

The standard deviation of the distribution σ is a measure of the uncertainty of the data provided by sensors. Durrant-Whyte [14] has used the summation of two Gaussian distributions to model uncertainty in the sensor measurement. Researchers have developed a few other

methods [15] to iteratively update the parameters of the distribution.

3.1 Estimation of Sensor Model Parameters

Maximum Likelihood (ML) method is the procedure for finding the value of one or more parameters for a given statistical data which maximizes the known likelihood distribution. If Gaussian distribution is considered, the distribution representing the sensor model is given by:

$$p_{D_i}\left(z_i \mid \sigma, x_i\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left\{\frac{-(z_i - x_i)^2}{2\sigma^2}\right\}} \tag{5}$$

where the event D_i represents the data (z_i, x_i), and $\Theta = \sigma$ is the parameter to be estimated. The likelihood function is the joint probability of the data given by:

$$L(\Theta) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\left\{\frac{-(z_i - x_i)^2}{2\sigma^2}\right\}} = \frac{1}{\sigma^n (2\pi)^{n/2}} e^{\left\{\frac{-\sum_{i=1}^{n}(z_i - x_i)^2}{2\sigma^2}\right\}} \tag{6}$$

and the parameter σ can be estimated via *ML* method by maximizing $L(\Theta)$ given by (6).

Most of the previous research on sensor fusion was based on the development of rigid sensor models. In practice, the performance of sensors or any source of information depends upon several factors, for example the environmental conditions under which the measurements were made, and the performance of estimation/calibration algorithm. Establishing dependence of a sensor’s performance on various parameters of environment and other signal/feature extraction algorithms is not a trivial task. Statistical techniques such as correlation analysis can be used to determine the manner in which these factors affect the sensor’s output. Selecting the factors that can possibly affect the sensor output is difficult, and is mostly based on heuristics. Many feature extraction algorithms include goodness-of-fit function that can be investigated to observe the correlation with sensor output.

After the factor which affects the sensor’s performance has been selected, the next challenge is to establish the correspondence between the factor and the uncertainty in the sensor’s output. Statistical system identification, regression analysis, or any mapping algorithm can be investigated to establish the correspondence. It might be difficult, if not impossible, to obtain the mathematical relation, and in the absence of such mathematical relation, model based statistical approach would be difficult to use. In this research project, the universal approximation capabilities of neural networks have been used to establish this correspondence.

3.2 Neural Network based Sensor Models

A neural network (NN) [16-17] is an information-processing paradigm inspired by the way in

which the heavily interconnected, parallel structure of the human brain processes information. They are often effective for solving complex problems that do not have an analytical solution or for which an analytical solution is too difficult to be found. Currently, they are being applied in many real world problems [18]. Three-layered NNs (i.e., one input layer, one output layer and one hidden layer), with hidden layer having sufficient nodes and a sigmoid transfer function, and linear transfer function in input and output layer [19-21] are considered to be universal approximators. In this research project, a three-layered NN has been formulated to obtain the correspondence between the standard deviation σ of Gaussian distribution and the parameter which affects sensor's performance. The output of a typical three-layered NN is given by:

$$O_i = f\left[\sum_{j=1}^{l}\left\{w_{ji}\, f\left(\sum_{k=1}^{n} w_{kj} I_k\right) + b_{1j}\right\} + b_{2i}\right] \quad i = 1, 2, \ldots \text{m} \tag{7}$$

where '*l*' is total number of nodes in the hidden layer, '*n*' is total number of inputs, '*m*' is total number of outputs, '*w*' is the weight and '*b*' is the bias of the network. The function *f(x)* is the activation function associated with nodes. The input to the neural network is the parameter vector I_k, and the output is the σ representing the standard deviation of the Gaussian distribution. Hence,

$$\sigma = f\left[\sum_{j=1}^{l}\left\{w_{ji}\, f\left(\sum_{k=1}^{n} w_{kj} I_k\right) + b_{1j}\right\} + b_{2i}\right] \tag{8}$$

or

$$\sigma = NNET(I, W, B) \tag{9}$$

where I is the vector representing input parameters, W is the weight matrix, and B is the bias matrix. Backpropagation (BP), based on gradient descent technique, is one of the most popular methods for training neural networks which establishes a particular set of weights obtained by adjusting the weights based on the errors between the actual and target output signals. For the neural network considered for the system in this research, however, the target data for σ is unknown, and cannot be obtained directly from experiments. Here, the neural network is trained in a novel manner from the signals obtained from *Maximum Likelihood* parameter estimation approach. Likelihood function that needs to be maximized is given by (6), in which parameter σ is represented by a neural network function given by (8) or (9). Hence, the likelihood function that needs to be maximized by choosing appropriate weights and biases of the neural network is given by:

$$L(W, B) = \frac{1}{\left[NNET(I, W, B)\right]^n (2\pi)^{n/2}}\, e^{\left\{\frac{-\sum_{i=1}^{n}(z_i - x_i)^2}{2[NNET(I, W, B)]^2}\right\}} \tag{10}$$

The weights and biases can be found by the gradient descent method or via evolutionary strategies [22]. The above method has been used to obtain models of infra-red proximity sensor

and vision sensors in stereo configuration.

3.3 Modeling of Stereo Vision Sensors

One of the most important components of stereo vision algorithm is stereo matching [23] which involves finding out the location of the point in right image plane corresponding to a point in the left image plane. The relative displacement of these two points, called disparity, is used to estimate the three-dimensional position of the point. The accuracy with which stereo vision sensors would be able to specify three-dimensional positional information about a point depends on how precisely the stereo vision algorithm is able to find the match of the point. The correlation score [24] of the matched points, which measures the correlation between two template windows from left and right images, is a measure of "goodness-of-match" of the two points. This score for template of size (2n+1) x (2m+1) is given by:

$$Score(P_L, P_R) = \frac{\sum_{i=-n}^{n}\sum_{j=-m}^{m}\left[I_L(u_L+i, v_L+j) - \overline{I_L(u_L, v_L)}\right]\times\left[I_R(u_R+i, v_R+j) - \overline{I_R(u_R, v_R)}\right]}{(2n+1)(2m+1)\sqrt{\sigma^2(I_L)\times\sigma^2(I_R)}} \qquad (11)$$

where I_L is intensity matrix of left image, I_R is intensity matrix of right image, and $\overline{I_k(u_k, v_k)}$ (*k=L, R)* is the average value of intensity, $\sigma(I_k)$ is the standard deviation of image I_k in the neighborhood of (2n+1) x (2m+1) of (*u*, v). The score ranges from -1 to +1, -1 representing not similar at all, and +1 representing most similar. The method formulated in the previous section has been used to develop a model for the stereo vision sensors that could take into account the performance of the stereo matching algorithm.

An experiment was carried out in the RAMA Laboratory, wherein a set of fifty data points consisting of 3-D location of point in world coordinate system obtained via stereo vision sensors (via transformation as discussed in reference [23]), correlation score for that point (given by (11)), and the actual 3-D location of the point in world coordinate frame. The value of correlation between the correlation score of stereo match of two image points and the error associated with that point in 3-D coordinates was found to be -0.3780, -0.2131, and -0.2856 respectively in X, Y, and Z direction. The error represents the absolute error between the actual 3-D location of a point and that obtained from the stereo vision. A negative correlation value represents that when correlation score is large, the error is small, which logically follows from the fact that larger correlation score means better stereo match and better estimation of 3-D positional information.

The strategy described in the previous section was used to develop a Gaussian model of the sensor. In this model the standard deviation of the distribution, which represents the uncertainty of the data, is dependent on the correlation score for the specific point. This dependence was modeled with the help of a neural network with five nodes in the hidden layer. This neural network takes correlation score as input, and outputs the value of standard deviation (sigma) for that particular correlation score. In order to ensure global convergence, the neural network was

trained via Genetic Algorithm. The sensor model obtained from this approach showed the intuitive trend that as the correlation score increases, i.e., as the stereo match gets better, the standard deviation decreases. Lesser value of standard deviations implies that the positional information obtained from stereo vision is less uncertain, and hence the degree of belief in the sensor output is more.

In order to investigate if anything was gained by carrying out the modeling with the help of neural network trained by maximum likelihood signal, the standard deviation was obtained by maximizing the likelihood function given by (6). This provided a constant value of standard deviation (sigma), representing rigid sensor model, for X, Y, and the Z directions for the same set of data. The NN based modeling that incorporated correlation score of stereo matching was able to improve the likelihood function by a factor 1.213, 1.1021, and 1.1288 in X, Y and Z directions respectively as compared to rigid sensor modeling approach. Hence, NN based modeling method promises to provide statistically more optimal and accurate results.

3.4 Modeling of Proximity Sensors

The output of the IR proximity sensor is an analog voltage which is indicative of the distance of the object detected by the sensor. In order to calibrate the IR sensor readings, 750 data points, comprising of sensor output values in volts and actual distance to the sensed object in mm, were taken. The calibration was obtained via neural network which accepted sensor output in volts as input and provided distance in mm as output. The neural network was trained from 750 data points obtained, and tested with the help of a separate set consisting of 700 data points. Fig. 2 shows the plot of neural network calibration, and test data. The plot shows sensor output (which is input to neural network) along Y axis, versus distance (which is output of the neural network) along X axis.

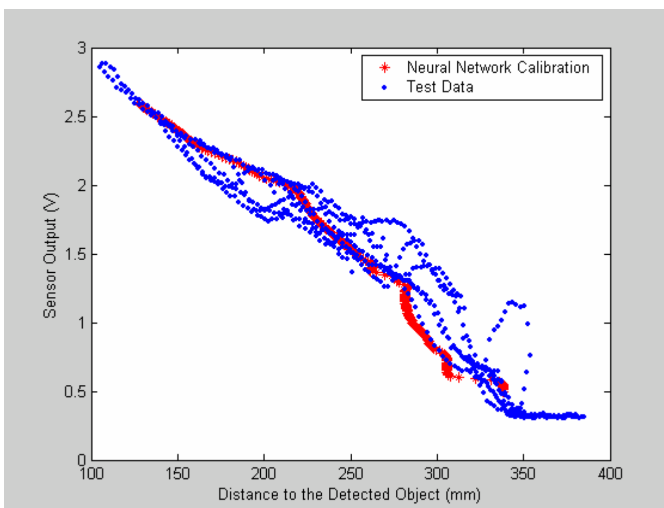

Fig. 2. Neural Network Calibration of IR Proximity Sensor

Once the calibration of the IR sensor was achieved, the next step was to develop the probabilistic model of the sensor based on the theory presented in the previous section. In order

to develop the sensor model, the most important aspect is to determine the factors that can possibly affect the performance of the sensor. From Fig. 2, where round dots show the test data obtained from experiments, it can be easily seen that the uncertainty in data increases as the distance to the object increases. This can be observed from the horizontal spread in the data as the distance to object increases. A larger horizontal spread means that for the same value of sensor output, the discrepancy in the actual distance to the object is more. In this research project, the model of the IR sensor attempts to capture this dependence of uncertainty on the distance to the object. Since the output of the sensor is indicative of the distance, this investigation makes use of neural network technique outlined in previous section to capture the relationship between sensor's uncertainties and sensor output.

In the laboratory experiments, the Infra-Red sensor was mounted on the wrist of the robot so that it looked vertically down (negative Z direction in world coordinate frame). The IR sensor provided the information about the distance to the nearest object detected directly in front of the object. Information about the position of end effector of the robot was obtained from the encoders of the robot. Hence, IR sensor can be effectively used in conjunction with robot encoders to provide 3-D information about any object. Similar to the case of vision sensor, the model of the IR sensor has been obtained in all three X, Y, and Z directions. The correlation value of the sensor output of a point and the error associated with that point was found to be -0.3078, -0.3211, and -0.2744 respectively in X, Y, and Z directions. The error represents the absolute error between the actual 3-D location of a point and that obtained from the infra-red sensor. A negative correlation value represents that when sensor reading is smaller (i.e., distance to the detected object is larger), the error is larger, which follows from large horizontal spread in Fig. 2 when the sensor reading is smaller. Neural network based modeling technique was used to obtain the senor model which represented the dependence of sensor's performance and inherent uncertainty on the distance to the detected object. The variation of standard deviation of the Gaussian sensor model obtained from this approach showed a decrease when the sensor output increased which implies that when the distance to the object decreases (i.e. sensor's output is larger) the standard deviation becomes smaller, and the sensor's measurement becomes less uncertain. This is also confirmed by Fig. 2 as well as the negative correlation value obtained above.

An analysis, similar to the one performed for vision sensors, was carried out to study the advantage of the proposed approach over the constant standard deviation (rigid sensor model) obtained by maximizing (6). The proposed NN based modeling was able to improve the likelihood function by a factor of 3.6690, 27.3505, and 2.3414×10^4 in X, Y, and Z directions respectively as compared to that obtained via rigid sensor modeling. Larger values of likelihood functions for modeling based on neural network method, similar to the case of stereo vision, reaffirms the fact that the proposed neural network based sensor modeling technique was more optimal and accurate.

4. SENSOR FUSION IN OCCUPANCY GRIDS

The occupancy grid [25-28] is a multi-dimensional field (usually of dimension two or three)

where each cell (or unit of the grid) stores or represents the probabilistic estimate of the state of spatial occupancy. Occupancy grids are one of the most common low-level models of an environment, which provide an excellent framework for robust fusion of uncertain and noisy data. If the state variable (occupancy, in this case) associated with a cell, C_i, is denoted by $s(C_i)$, then the occupancy probability $P[s(C_i)]$ represents the probabilistic estimate of occupancy of that particular cell. If $P[s(C_i)=occ]\approx 0$, then the cell is assumed to be empty, while, if $P[s(C_i)=occ]\approx 1$, then the cell is assumed to be occupied.

If a single sensor is used to obtain the occupancy grid, Bayes' Theorem can be used in the following manner to determine the state of the cell:

$$P[s(C_i)=occ \mid z]=\frac{p[z \mid s(C_i)=occ]P[s(C_i)=occ]}{\sum_{s(C_i)} p[z \mid s(C_i)]P[s(C_i)]} \tag{12}$$

where z is the sensor measurement. The probability density function (p.d.f.) $p[z|s(C_i)=occ]$ is dependent on the sensor characteristics and is called the sensor model. The probability $P[s(C_i)=occ]$ is called prior probability mass function and specifies the information made available prior to any observation.

The sensor models obtained in Section III represent the probability of a sensor providing the position z when actual position of the object is x. These models relate the probability of occurrence of an object to the distance from the measurement (in one dimension), given that the object was detected by the sensor. These models (likelihood functions) alongwith any prior information is used to obtain the posterior distribution $p(x|z)$ which represents uncertainty in the positional information supplied by the sensors. In this research, an occupancy profile has been recreated with the help of three-dimensional occupancy grids. Each cell of the grid represents a three-dimensional space in the world coordinate system. The occupancy profile of the workspace is obtained by determining whether the cells of the grids are occupied or not. A cell is said to be occupied if the spatial region represented by this cell has at least one point (detected by sensors) within it. Fig. 3 shows a situation in which two points contribute to the probability of a cell (in one dimension) being occupied. The probability that point A lies within the cell C_i, represented by the space from x_{i-1} to x_i, can be obtained from the equation:

$$P(A \in C_i)=\int_{x_{i-1}}^{x_i} p(x|z_A)dx \tag{13}$$

where z_A is the sensor reading, and $p(x|z_A)$ is the posterior distribution. Similarly, for point B:

$$P(B \in C_i)=\int_{x_{i-1}}^{x_i} p(x|z_B)dx \tag{14}$$

The probability that either A or B is in the cell C_i is given by:

$$P[(A \in C_i) U (B \in C_i)] = P(A \in C_i) + P(B \in C_i) - P(A \in C_i) P(B \in C_i) \tag{15}$$

The above equation makes an assumption that detection of points *A* and *B* are independent processes. The above approach can be used to determine the probability of cell being occupied by at least one point among any number of points detected in the vicinity, and is used to obtain individual occupancy grids from IR proximity sensor and stereo vision.

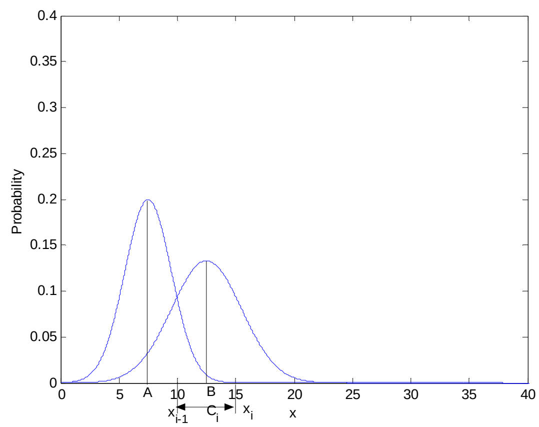

Fig. 3. Two Points Contributing to Occupancy of a Cell

4.1 Fusion of Two Occupancy Grids

The fusion of grids obtained from individual sensors is carried out by determining probabilistic estimates at the cellular level. Most of the previous research work makes use of probabilistic evidence combination formula based on Independent Opinion Pool [29] and use of maximum entropy priors [25, 27]. This method is applied at each cell to fuse two probabilistic estimates, P_1 and P_2, associated with the two grids obtained from two sensors. The mathematical expression representing the fusion is given by:

$$P[s(C_i) = occ \mid P_1, P_2] = \frac{P_1 P_2}{P_1 P_2 + (1 - P_1)(1 - P_2)} \tag{16}$$

There are certain problems associated with the use of this equation to fuse two estimates. For example, this equation fails to provide a reasonable estimate if the two sensors have completely contradictory measurements. Moreover, this equation puts equal amount of belief in the two measurements, and would provide incorrect estimates if the reliability/accuracy of the two sensors that are being fused vary by a large amount. In this research, for example, stereo vision sensors are fused with infra-red sensors. The stereo vision sensors provide very accurate measurements in X direction, while the infra-red sensor does not. It would be unreasonable to provide equal weight to the probabilistic estimates obtained from these two sensors. In this research project, the two estimates are fused in a Bayesian framework taking into account the

certainty and reliability of the sensors. As described in previous sections, sensor models are given by the following Gaussian likelihood function:

$$p(z_k \mid x) = \frac{1}{\sigma_k \sqrt{2\pi}} e^{\left\{ \frac{-(x - z_k)^2}{2\sigma_k^2} \right\}} \qquad k=1,2 \tag{17}$$

where $k=1$ represents stereo vision, and $k=2$ represents IR sensor.

Then, from Bayes' Theorem the fused *MAP* estimate is given by:

$\hat{x}_{MAP} = \arg\max\left[p(z_1|x)p(z_2|x)\right]$ or

$$\hat{x}_{MAP} = \arg\max\left[\frac{1}{\sigma_1 \sigma_2 2\pi} e^{-\left\{ \frac{(x-z_1)^2}{2\sigma_1^2} + \frac{(x-z_2)^2}{2\sigma_2^2} \right\}} \right] \tag{18}$$

which gives:

$$\hat{x}_{MAP} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} z_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} z_2 = \frac{1}{r^2 + 1} z_1 + \frac{1}{1 + \frac{1}{r^2}} z_2 \tag{19}$$

where, $r = \sigma_1 / \sigma_2$ is the ratio of standard deviations. Hence, if there is no prior information available about the quantity to be estimated, the Bayesian approach for fusion of the two sensor estimates results in a weighted average dictated by the ratio of standard deviations. If two Gaussian distributions (given by the two sensor model's pdfs) are fused, then the posterior distribution is jointly Gaussian with a mean given by (19) and the standard deviation given by:

$$(\sigma')^2 = \left[(\sigma_1)^{-2} + (\sigma_2)^{-2}\right]^{-1} \tag{20}$$

Fig. 4 shows the two distributions that get fused to give the posterior distribution. The simple product of distribution is also shown in the figure. It may be noted from the figure that the standard deviation of fused distribution is smaller representing lesser uncertainty in fused estimates.

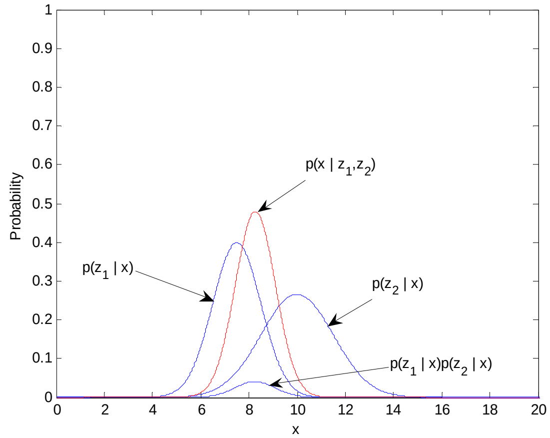

Fig. 4. Fusion of Two Gaussian Distributions

5. EXPERIMENTAL RESULTS AND DISCUSSION

The theories developed in the previous sections were validated with the help of experiments performed in the Robotics and Manufacturing Automation (RAMA) Laboratory at Duke University. A cylindrical object was placed on the work-table. Fig. 5 shows the images of the work-table obtained from the stereo cameras. Fig. 6(a) shows the actual occupancy grid of the workspace. This was obtained based on the geometric dimensions of the object and its location in the workspace. For the occupancy grid developed in this research, each grid is of size 5mm X 5mm X 5mm.

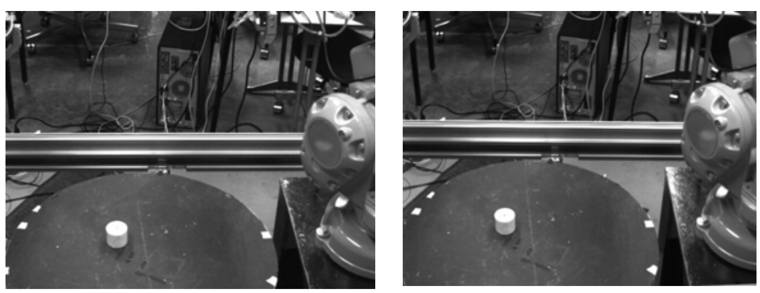

Fig. 5. Images of the Worktable Obtained from Left and the Right Camera

Once the data points were obtained from the IR proximity sensor and stereo vision sensors, Equations (13) to (15) were applied to find the probabilistic estimates of the occupancy of cells of the grid, and two separate occupancy grids obtained from the IR and the stereo vision were developed. If the probability of the cell being occupied was $P[s(C_i) = occ] > 0.5$, the cell was assumed to be occupied. The occupancy grid obtained from fusion of the two grids via Bayesian approach described in previous section is shown in Fig. 6(b).

The occupancy grids obtained above were based on the sensor models derived from the proposed neural network based learning scheme. If the sensor models are assumed to be rigid

(standard deviation does not depend on any parameter and remains constant for all data points), a different set of occupancy grids is obtained. Fig. 6(c) shows the occupancy grid fused from these two grids. In order to carry out the comparison, fused occupancy grid was also obtained from the independent opinion pool method given by (16), and the fused grid is shown in Fig. 6(d)

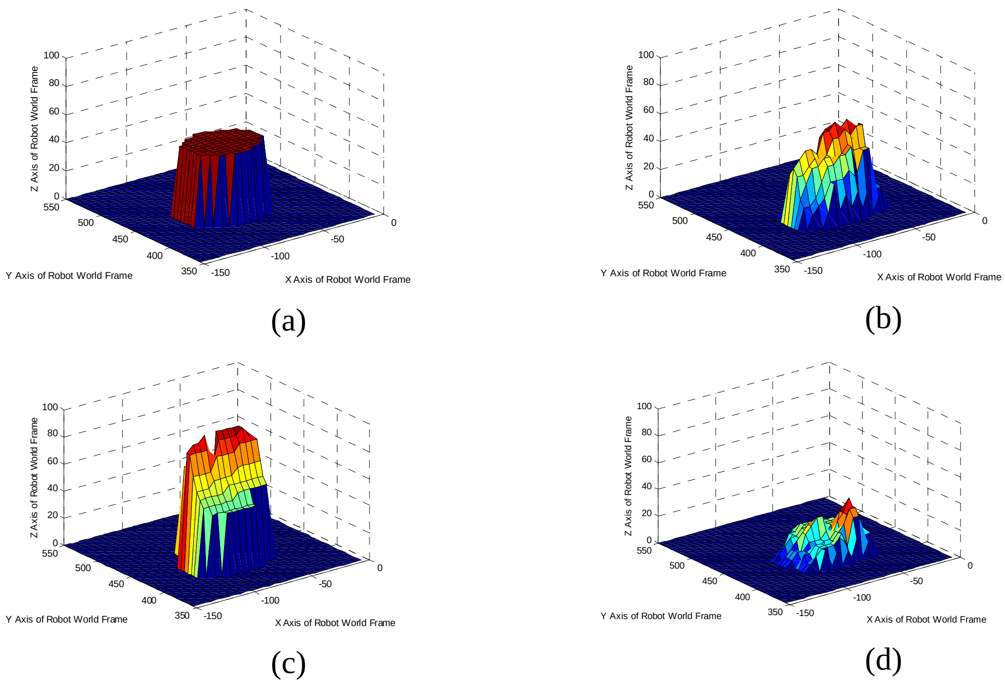

Fig. 6. Occupancy Grids a) Actual Grid, b) Fused Grid (Proposed NN based Approach), c) Fused Grid (Rigid Sensor Model Approach), d) Fused Grid (Independent Opinion Pool Method)

To facilitate the comparison of performance of the fusion process via different algorithms, a measure of error was formulated which is given by the following equation:

$$Error = \sum_{C_i} \left[\left| s(C_i) \right|_{actual} - \left| s(C_i) \right|_{sensor} \right]^2 \qquad (29)$$

where $\left| s(C_i) \right|_{actual}$ is the actual state of the cell, and $\left| s(C_i) \right|_{sensor}$ is the state of the cell obtained from the sensor and/or fusion process. The state of the cell is either 1 (for occupied) or 0 (for empty). The value of this error associated with the occupancy grids obtained from raw sensor measurements are 1062 and 1279 respectively for IR proximity sensor and stereo vision. Table I provides the error value associated with the occupancy grid obtained from the fusion process described above. The table compares the error value obtained via the two approaches. The first approach is based on the proposed neural network oriented sensor modeling scheme, and the second approach is based on the rigid sensor model.

TABLE I
ERROR ASSOCIATED WITH OCCUPANCY GRIDS OBTAINED FROM FUSION PROCESS

NN Based Modeling Approach			Rigid Sensor Model Approach		
IR Sensor	Stereo Vision	Fused	IR Sensor	Stereo Vision	Fused
1070	1160	956	1070	1279	1214

From the figures as well as from the table of results, it is evident that the sensor modeling and fusion scheme presented in this research report have been able to reduce the uncertainty inherent in individual sensors. The fusion process derived from proposed neural network based sensor modeling scheme has been able to reduce the error associated with the occupancy grid by over 25% with respect to raw stereo vision sensor measurements, and by approximately 10% with respect to raw IR proximity sensor measurements. On the other hand, the results from fusion based on rigid sensor model scheme are not particularly impressive. Although the occupancy grid obtained from this method had a reduction of error by 5% with respect to raw stereo vision sensor measurement, the fused grid had an increase in error by over 14% with respect to raw proximity sensor measurements. The major reason for its poor performance is that this procedure assumes equal uncertainty (given by constant standard deviation of sensor model) associated with all data points. This assumption leads to erroneous calculation of probabilistic estimates as shown by an increase in the error value. The error associated with the occupancy grid obtained from independent opinion pool method was found to be 1063. Hence, it shows no improvement when compared to raw proximity sensor measurement, and shows an improvement of approximately 17% with respect to stereo vision sensor measurements. This method tends to yield high probabilistic evidence when both the sensors are in agreement, and gives low probabilistic evidence when at least one sensor provides lower value compared to the other. Independent opinion pool method indicates equal belief in the probabilistic evidence provided by two different sources. For fusion problem investigated in this research that involved sensors which have large difference in reliability/accuracy of their measurement, this method is not suitable.

6. CONCLUSION

This research project proposes a novel technique that utilizes learning and optimizing capability of neural networks to obtain a sensor model that automatically learns some of the key statistical relations from the data. The research results present a method of obtaining occupancy profile of the environment based on a three-dimensional occupancy grid framework, and present a novel method for obtaining probabilistic estimates for the occupancy of cells from two different kind of sensory sources. The fusion of information from sensors has been carried out under Bayesian framework, and its performance has been compared with respect to other strategies. It was seen that the fusion process carried out under proposed sensor modeling strategy based on neural network had a superior performance than that of fusion process based on rigid sensor model and that of fusion based on independent opinion pool method.

8. PUBLICATIONS RESULTING FROM THIS PROJECT

- Kumar, M., Garg, D., and Zachery, R., "Multi-Sensor Fusion Strategy to Obtain 3-D Occupancy Profile", accepted for publication in the *Proceedings of the 31st Annual Conference of the IEEE Industrial Electronics Society (IECON)*, Raleigh, NC, November 2005.

- Kumar, M., Garg, D., and Zachery, R., "Intelligent Sensor Modeling and Data Fusion via Neural Network and Maximum Likelihood Estimation", accepted for publication in the *Proceedings of the ASME International Mechanical Engineering Congress and Exposition*, Orlando, FL, November 2005.

- Kumar, M., Garg, D., and Zachery, R., "Intelligent Sensor Uncertainty Modeling Techniques and Data Fusion", submitted to *IEEE Transactions on Mechatronics*, 2005.

REFERENCES

[1] Shafer, G., *A Mathematical Theory of Evidence*, Princeton, NJ: Princeton University Press, 1976.

[2] Demspter, A. P., "A Generalization of Bayesian Inference", *Journal of Royal Statistical Society, B*, Vol. 30, No. 2, 1968, pp. 205 – 247.

[3] Yager, R. R. and Zadeh, L. A., (Eds.), *An Introduction to Fuzzy Logic Applications in Intelligent Systems*, Kluwer Academic Publishers, 1991.

[4] Klir, J. K. and Yuan, B., *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Upper Saddle River, NJ: Prentice-Hall, 1995.

[5] McKendall, R. and Mintz, M., "Data Fusion Techniques Using Robust Statistics", *Data Fusion in Robotics and Machine Intelligence*, Abidi, M. A. and Gonzalez, R. A. (eds.), Academic Press, 1992.

[6] Maybeck, P. S., *Stochastic Models, Estimation and Control*, Volume 1, Academic Press, Inc., 1979.

[7] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems", *Transactions of the ASME-Journal of Basic Engineering*, Vol. 82, Ser. D, 1960, pp. 35 – 45.

[8] Sasiadek, J. Z., "Sensor Fusion", *Annual Reviews in Control*, Vol. 26, 2002, pp. 203-228.

[9] Luo, R. and Su, K., "A Review of High-Level Multisensor Fusion: Approaches and Applications", *Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems,* 1999, pp. 25-31.

[10]Press, S. J., *Bayesian Statistics: Principles, Models and Applications*, John Wiley and Sons, 1989.

[11]Clark, J. J., and Yuille, A. L, *Data Fusion for Sensory Information Processing Systems*, Kluwer Academic Publications, 1990.

[12]Bayes T., "An Essay Towards Solving a Problem in Doctrine of Chances", *Philosophical Transactions*, Vol. 53, 1763, pp. 370 – 418.

[13]Manyika, J. and Durrant-Whyte, H., Data Fusion and Sensor Management: a Decentralized Information-Theoretic Approach, Ellis Howard Limited, 1994.

[14]Durrant-Whyte, H.F., Integration, Coordination and Control of Multi-Sensor Robot Systems, Kluwer Academic Publishers, Norwell, MA, 1988.

[15]Porrill, J., "Optimal Combination and Constraints for Geometrical Sensor Data", *The International Journal of Robotics Research*, 1988, pp. 66 – 77.

[16]Rumelhart, D. E. and McClelland, J. L., *Explorations in Parallel Distributed Processing: a Handbook of Models, Programs, and Exercises* (MIT Press), 1988.

[17]Haykin, S. S., Neural Networks: A Comprehensive Foundation (Prentice Hall Press), 1998.

[18]Garg, D., Ananthraman, S., and Prabhu, S., "Neural Network Applications", *Wiley Encyclopedia of Electrical and Electronic Engineering*, John G. Webster (ed.), Vol. 14 (John Wiley, New York, NY), 1999, pp. 255 – 265.

[19]Kumar, M. and Garg, D., "Neuro-Fuzzy Controller Applied to Multiple Robot Cooperative Control", *International Journal of Industrial Robot*, Vol. 32, Issue 3, 2005, pp. 234 – 239.

[20]Hornik, K., Stinchombe, M., and White, H., "Multilayer Feedforward Network are Universal Approximators", *Neural Networks*, Vol. 2, 1989, pp. 359 – 366.

[21]Irie, B. and Miyake, S., "Capabilities of Three-Layered Perceptrons", *Proceedings of IEEE International Conference on Neural Networks*, California, USA, 24-27 July, 1988, pp. 641 – 648.

[22]Garg, D. and Kumar, M., "Optimization Techniques Applied To Multiple Manipulators for Path Planning and Torque Minimization", *Journal for Engineering Applications of Artificial Intelligence*, vol. 15, 2002, pp. 241-252.

[23]Garg, D. and Kumar, M., "Object Classification via Stereo Vision in a Flexible Manufacturing Work Cell", *Proceedings of the 10th International Conference on Mechatronics and Machine Vision in Practice*, Perth, Western Australia, December 9-11, 2003.

[24]Zhang, Z., Deriche, R., Faugeras, O., Luong, Q.-T., “A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry”, *Artificial Intelligence Journal*, Vol.78, October 1995, pp. 87 – 119.

[25]Elfes, A., “Multi-Source Spatial Data Fusion Using Bayesian Reasoning”, *Data Fusion in Robotics and Machine Intelligence*, Abidi, M. A. and Gonzalez, R. A. (eds.), Academic Press, 1992.

[26]Kumar, M. and Garg, D., “Intelligent Multi Sensor Fusion Techniques in Flexible Manufacturing Workcells”, *Proceedings of American Control Conference*, 2004, Boston, MA, pp. 5375 – 5380.

[27]HoseinNezhad, R., Moshiri, B., Asharif, M.R., “Sensor Fusion for Ultrasonic and Laser Arrays in Mobile Robotics: a Comparative Study of Fuzzy, Dempster and Bayesian Approaches”, *Proceedings of IEEE Conference on Sensors*, Vol. 2, 12-14 June 2002, pp. 1682 – 1689.

[28]Kumar, M. and Garg, D., “Three-Dimensional Occupancy Grid with the Use of Vision and Proximity Sensors in a Robotic Workcell”, Paper Number IMECE2004-59593, *Proceedings of the ASME International Mechanical Engineering Congress and Exposition*, Anaheim, CA, November 14-19, 2004, 8p.

[29]Berger, J. O., *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, Berlin, 1985.